



DIGITAL ACCESS AND PRESERVATION POLICY



Digital materials at the UAA/APU Archives and Special Collections are in a variety of formats. The Archives has in its holdings digitized photographs, either for preservation purposes or reference, born digital content, digitized copies of audio, video, and film, as well as photographs taken of materials not usually kept within the collections, i.e. awards and plaques. The Digital Access and Preservation Policy (DAPP) will outline, in further detail, the different types of digital content within our holdings, how each should be treated, and how each should be saved.

The Archives uses a Network Attached Storage (NAS) device, which has been separated into two main drives: Drive 1 and Drive 2. Drive 1 is the only one that is backed-up to the cloud service provider, Backblaze. Once materials are saved onto Drive 1, they should not be accessed. Drive 1 is to act as a dark archive. Therefore, we created Drive 2 where archivists can retrieve the digital contents for reference and use purposes. There is also Drive 3, which should be mapped separately to each computer. Drive 3 is the drive our users will use to access digital content while they are in the Research Room. Drive 3 on the researcher computer in the Research Room is a read-only drive, which means that users cannot make changes to any items within this drive.

Description of materials on each drive

Drive 1 on NAS

High resolution copies of everything as well as born digital files. This includes the high resolution copies of the digitized audio, video, and film as well as images, documents, and all born digital files provided by donors.

In summary, everything that was digitized high resolution (audio, video, film, and photographs), all of our born digital materials, digitized copies of documents, and digitized copies of photographs and documents that were either removed from collections and destroyed or are inaccessible for preservation purposes. This also includes awards, plaques, and certificates that were removed from collections.

Drive 2 on NAS

Only the access copies of the digitized copies of our audio, video, and film.

Born digital images and documents and files provided by the donors. This also includes the digitized copies of original documents that have been destroyed or inaccessible for preservation purposes (copies of photographs and documents that were either removed from collections and destroyed or are inaccessible for preservation purposes). This also includes awards, plaques, and certificates that were removed from collections.

High resolution copies of digitized photographs that are in our collections.

In summary, this drive is a copy of the Master drive, with the exception of the audio, video, and film that we had digitized and the original item is still in the collection.

Drive 3 (mapped separately and not on the NAS)

All low resolution born digital and nitrate images.

Copies of born digital materials, donor provided materials, and digitized copies of documents that were removed from collections and destroyed for preservation purposes. Some of these materials will be in all three drives, such as pdfs, video, and photographs provided by donors (for example Rage City Roller Girls). This also includes awards, plaques, and certificates that were removed from collections.

Access copies of the digitized audio, video, and film in collections.

In summary, low resolution copies of born digital photographs, access copies of the audio, video, and film, and all the files that only exist in an electronic format. This will not include digitized copies of the high resolution photographs from our collections that are saved on the NAS drives. (For example, the digitized photographs in Fickett will not be converted to low resolution and will not be put in this folder).

Types of digital materials

Photographs digitized for preservation purposes: These include the high-resolution copies of digitized photographs that are in danger of deteriorating based on their age or make, specifically nitrate negatives or photographs which are oxidizing. These materials should be scanned at an appropriate resolution so they will not need to be scanned again if requested by a user. (It is recommended these materials should be scanned at 2400 dpi, with 10 inches on the long edge). Low resolution copies of these items should also be created to make them more easily accessible in the Drive 3.

Where to save these items: The high resolution copies should be saved in Drives 1 and 2. It is recommended to save the low resolution copies in all of the drives (Drives 1, 2, and 3), as they can take a while to create, especially if there is a lot in one collection. All items within each drive should be saved in the appropriate collection folder, under the sub-folder “Digitized_for_preservation”.

Photographs digitized for reference: These are digitized copies of photographs within our collections where a high-resolution digital copy has been requested by a researcher.

Where to save these items: These should only be saved within the Drives 1 and 2. They should be saved in the appropriate collection folder, under the sub-folder “Digitized_for_reference”.

Documents digitized for reference: These should only be saved when the whole folder has been scanned. This also includes anything that has been scanned for reference purposes that had to be redacted. For naming redacted files please put “-redeacted” in the file name. For example: “uaa-hmc-0001-b2-f03-redeacted”.

Where to save these items: These should be saved in the Drives 1 and 2. They should be saved in the appropriate collection folder, under the sub-folder “Digitized_for_reference”.

Audio, video, and film digitized for preservation: These should include both the high resolution and low resolution digitized copies of the A/V materials. In the case of video and film, high resolution should be AVI files, and low resolution should be MP4 files. For audio files, high resolution should be WAV master files and 192 kbps MP3 files.

Where to save these items: Both the high resolution and low resolution digital copies should be saved in Drive 1, and only the low resolution copies should be saved to the Drive 2 and Drive 3. They should be saved in the appropriate collection folder, under the sub-folder “Digitized_for_preservation”.

Born digital documents: Born digital documents that were digitized by the donor (whether or not they donated the original items as well), and items created electronically.

Where to save these items: Drives 1 and 2 should include all original items, as well as Drive 3. They should be saved in the appropriate collection folder, under the sub-folder “Born_digital”.

Note: It is not usually necessary to lower the resolution or compress these files in order for them to be browsable in Drive 3.

Born digital photographs: Born digital photographs include photographs digitized by the donor or created electronically.

Where to save these items: The Drives 1 and 2 should include all original items. They should be saved in the appropriate collection folder, under the sub-folder “Born_digital”. If necessary, depending on the file type and size of the original photographs, access copies should also be created of these items and saved with the original items in the Drives 1 and 2, within the “Born_digital” folder. And only the lower resolution access copies should be saved to Drive 3. This will make the files easily browsable for those using the Research Room computer. For example: The Anne Nevaldine papers included 30,000 digital photographs, a majority of which were Canon Raw Files (CR2). These are not easily browsable as they are fairly large files and take a long time to load. Low resolution JPGs were created for all of the digital photographs within the collection and the original file structure was maintained. All of the photographs are under the “Born_digital” folder of the Nevaldine collection, however, there are two subfolders, one with the original file types, and one with the JPGs. In the case of the JPGs original folder structure was maintained, so it is essentially a duplicative folder of all the photographs, but with JPGs instead of CR2 files.

Born digital audio and video: Audio and video files that were created or digitized by the donor (whether or not they donated the original items as well).

Where to save these items: Drives 1, 2, and 3 drives should include all original items. They should be saved in the appropriate collection folder, under the sub-folder “Born_digital”. If necessary (depending on format and file size) access copies should also be created of these items and also saved along with the original items, but within a separated folder.

Awards, certificates, and plaques: These are usually photographs taken by the Archives or donors of the awards, plaques, and certificates within a collection. The Archives does not typically keep the original items because they are hard to preserve, have limited research utility, and can take up a lot of space within the vault.

Where to save these files: The photographs of these items should be saved within Drives 1, 2, and 3. They should be saved in the appropriate collection folder, under the sub-folder “Digitized_for_preservation”.

Setting up the drives

All digital materials should be saved within their collections using the collection number as the main folder (i.e. HMC-0001), and within one of the three main sub-folders reflecting the type of digital records. The only sub-folders that should be reflected within each collection are:

- Born_digital
- Digitized_for_preservation
- Digitized_for_reference*

*Note, Drive 3 will not have a "Digitized_for_reference" folder. This folder should be in the Drives 1 and 2 drives only.

Examples:

Drive 1

- HMC-0001
 - Born_digital
 - Digitized_for_preservation
 - Digitized_for_reference

Drive 2

- HMC-0001
 - Born_digital
 - Digitized_for_preservation
 - Digitized_for_reference

Drive 3

- HMC-0001
 - Born_digital
 - Digitized_for_preservation

It is also important to appropriately organize all files within these folders. The organization is largely dependent on the collection. Born digital items, especially if given by a donor, should reflect the original arrangement and folder/file names. Items digitized for reference and preservation should be saved either under the document's identifier number, or in the case of a large number of files, the box number.

Some examples of folder structures include:

- HMC-0001 (Note: Drives 1 and 2 only)
 - Digitized_for_reference
 - Box21
 - Box22
 - Box23

- HMC-1234
 - Digitized_for_preservation
 - NitrateNegatives
 - UAA-HMC-1234-B12-F25
 - UAA-HMC-1234-B12-F26
 - UAA-HMC-1234-B12-F27
 - Album1
 - Album2
- HMC-1355 (Note: within Drive 1 only)
 - Digitized_for_preservation
 - AVI
 - UAA-HMC-1355-B1-item01
 - UAA-HMC-1355-B1-item02
 - UAA-HMC-1355-B1-item03
 - MP4
 - UAA-HMC-1355-B1-item01
 - UAA-HMC-1355-B1-item02
 - UAA-HMC-1355-B1-item03
- HMC-1355 (Note: within the Drives 2 and 3 only)
 - Digitized_for_preservation
 - MP4
 - UAA-HMC-1355-B1-item01
 - UAA-HMC-1355-B1-item02
 - UAA-HMC-1355-B1-item03

Online backup

The Drive 1 is backed up every evening to a cloud software backup provider. Archives and Special Collections does not have access to this service. Only the Library's PC and Network Administrator has the access. When changes are made to Drive 1, the backup does not read the changes, but creates duplicative folders (in the case of a name change), and does not remove a file or folder from the backup when it has been deleted from Drive 1. It is recommended to organize all files within a separate drive or external hard drive before saving to Drive 1. If any changes do need to be made to Drive 1, it is okay to make them, but be sure to inform the Network Administrator of the changes so he can make the same one within Backblaze. When emailing him, it is important to tell him what collection, what folder, and what document(s). For example:

“The following is a list of deletions for the NAS backup. In Drive 1, all of these will be within HMC-1180 under Digitized_for_preservation:

In the main folder, please remove the following folders and files:

‘Parker’

‘New folder’

‘AMIPA_uaa-hmc-1180-b42-f05’

‘ObjectInfo”

‘uaa-hmc-1180-b42-f06’

Within the folder ‘Box40’ please remove:

‘uaa-hmc-1180-b40-f34’”

Checksums

Initial checksums for materials being transferred to the NAS

All digital files that come to the Archives, whether they are born digital or digitized in-house or by a vendor (e.g. Scene Savers), need to have checksums run on them before and after we transfer them to the NAS. Use ExactFile to create a digest of files to be transferred, and save it to erecordschecks > InitialChecksums > [collection #]. If there is not already a folder for the collection that the material belongs to, create one.

Once the files have been transferred over to the NAS, use the Test Digest tab in ExactFile to test the copies of the files in Drives 1, 2, and 3 against the Digest you created. If the checksums match for all of the files, the process is complete. If there is an error, you may need to transfer the file(s) again. Remember that you need to let the Network Administrator know if you delete anything from Drive 1 so that he can delete it from the cloud storage service.